DOCUMENT RESUME

ED 240 009 SO 015 353

AUTHOR Wallen, Norman E.
TITLE What Is Wrong with Research in Social Studies?
PUB DATE Nov 83
NOTE 7p.; CUFA Presentation to the National Council for the Social Studies (San Francisco, CA, November 23, 1983).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cognitive Tests; *Educational Needs; *Educational Research; Elementary Secondary Education; Evaluation Methods; Measurement; Measurement Techniques; Reliability; Research Design; *Research Problems; *Research Utilization; *Social Studies; Statistics; Teacher Effectiveness; Validity
IDENTIFIERS Inference (Statistical); Replication

ABSTRACT

Social studies research has told us very little in the past 30 years. There are several reasons for this state of affairs in social studies. First, inferential statistics has become much too prominent in research thinking and practice. Replication should be used by researchers to build a useful body of knowledge. A second reason for lack of success is the continuing failure to be really concerned about measurements. Reliability and validity need to be addressed, and new tests for measuring higher level cognitive skills need to be developed. Examples from the Taba curriculum project and from a recent research report which appeared in the journal "Social Education" show the difficulties encountered in measuring students' abilities to make inferences and to explain and apply ideas and students' reasoning about basic economic concepts. Three mistakes made in both measurement attempts were: (1) looking too much to other disciplines for guidance in studying the educational process, (2) trying to study underlying psychological processes instead of important content, and (3) unnecessarily and somewhat inappropriately over-statisticizing. A more idiographic approach to measurement needs to be taken. (RM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED240009

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

✓ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Norman E. Wallen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

WHAT IS WRONG WITH RESEARCH IN SOCIAL STUDIES?

by
Norman E. Wallen
San Francisco State University

SO 015 353

CUFA Presentation to the National Council for the Social Studies
San Francisco, CA
November 23, 1983

NCSS 11/23/83

What Is Wrong With Research in Social Studies?

Norm Wallen

San Francisco State University

My point of departure for this paper is to address the question: "What is wrong with research in social studies?" This, of course, implies that something is wrong so, at the risk of belaboring the obvious, let me assert that research in social studies has told us very little in the past 30 years. Specifically, we have learned very little about our effectiveness beyond the trite observation that some approaches seem to work sometimes but not other times. We have little evidence to support our belief that we can, as teachers, do specific things to affect student values, knowledge, and cognitive skills. It is this, I submit, as much as reactionary, pecuniary politics that has reduced social studies to secondary status in our schools.

Why is this so? Let me offer a few possibilities. First, I think we have allowed inferential statistics to become much too prominent in our research teaching and practice. I say this with regret since my major emphasis in graduate school was statistics and I dearly love explaining the logic behind analysis of variance. Nonetheless, I find a certain lack of logic in requiring students to master the intricacies of testing for adherence to minor assumptions such as homogeneity of variance when the major assumption underlying all statistical inference, i.e., randomness of sampling, is virtually never met, and is in fact, glossed over lightly in most texts.

This is not to say that the intent of inferential statistics is misguided. It is, after all, an attempt to generalize research findings, an inevitable undertaking unless one is completely unintellectual. As an aside, I occassionally run across colleagues who insist they don't want to generalize; they are merely describing what happened. My question then is: "Why tell anyone then--much less publish your results?"

How then are we to build a body of knowledge which is useable? I am afraid we are stuck with replication. Replication actually has a very reputable history--mostly in the physical sciences. It is an interesting question as to why the behavioral sciences have chosen, in this regard, to emulate agricultural [redacted] research rather than physics and chemistry. Granted, replication is more difficult outside the laboratory, but not sufficiently so to account for its rarity. Somehow, it is not quite respectable. I once succeeded in getting a University department to accept replications as M.A. theses; during the next five years, no student ever took that alternative; they all wanted to express their creativity--I guess.

A second reason I suggest for our lack of success is our continuing failure to be really concerned about our measurements. How many research reports seriously address the reliability and validity of their measures? How many discuss the matter at all? Is it true that we simply cannot measure those variables most important to us-higher level cognitive skills,

knowledge, and values? I don't think so, though I think we have been kidding ourselves as to how it can be done. Isn't it time to admit that we really can't do the job with paper and pencil tests? Now that we know so much about what's wrong with tests like the Binét and Wechsler, isn't it time to follow their model of format and development but with a different set of tasks?

Thinking about this paper caused me to look again at my own principal experience along these lines. It occurred on the Taba Curriculum Project, which some of you may recall, We did our best to develop a series of paper and pencil tests to measure such things as ability to explain, make inferences, and apply ideas. It remains a major disappointment to me that we were not able to show much difference on these measures between students who were taught Taba for a year and those who were not (you will recall that the Eight Year Study had a similar problem).

There are, obviously, plenty of explanations: too short a time, poor implementation of the program, etc., but I think the main problem is something else. I think we were going about it the wrong way. By using a typical "across the board" standard set of items, we may have obscured what we wanted to measure. Suppose that a particular sixth grader had, at the beginning of school, no conception of interdependence among groups (a safe assumption since most adults don't seem to have it). It still seems probable to me that that student was much more likely to grasp and use the idea if he or she were in the Taba program--and that would be important. However, that change might well have gotten lost by being measured by only a few of our 65 items--which had to assess a variety of concepts, and by the nature of the items themselves. Item analysis wouldn't necessarily help because most students might not have changed on any given item. What I am arguing for is a more idiographic (remember Allport) approach to measurement. Suppose we, or the teacher, had spent more time at pre-testing finding out which concepts/ideas each student was weak in and re-assessing them, or at least those with highest priority, at post-test. We might have found that different children had learned a lot with regard to specific ideas, not necessarily all of them. By focusing on a generalized 'ability to apply ideas' we might have missed the more important finding that most students acquired a few very important ideas.

This approach isn't original, I'm afraid. To my knowledge, it has been suggested in the evaluation context by Kirasuk (1973) and is the method of choice in that area of Special Education dealing with severely handicapped students. Admittedly, the latter setting differs from social studies. It is one thing to focus on a severely handicapped child's ability to tie shoes and another to focus on a typical sixth graders understanding of interdependence. Closer to home, however, you may recognize the similarity of my proposal to the concept of prescriptive teaching propounded by Lindsley. I think one reason so few teachers can be found following Lindsley's model is that it is tedious and boring to continually monitor progress in fractions or spelling. This need not be so in social studies. Our content is, I believe, inherently interesting to both teachers and students, provided they are engaged in intellectual exploration of ideas. Why not allow teachers to monitor grasp of concepts and ideas through both anecdotal records and use of techniques such as interview and role playing?

An approach which is consistent with my argument is that of Freyberg and Osborn who have been conducting in-depth interveiws to determine the understanding which New Zealand students at various ages have of concepts in the physical and biological sciences. They find substantial evidence of misunderstanding. For example, many 10-12 year-olds believe the "gram" to be a measure of length. Others understand the concept of gravity to apply only to earth. Often these misunderstandings are masked by the complexity of communication and come to light only after skillful questioning. I suspect someone has done a similar analysis of concepts in social studies but I am not aware of it. Nevertheless, I am certain that the misunderstanding or outright incomprehension of many of our central concepts is rampant. Parenthetically, I am also convinced from many conversations that many of my colleagues in education and in psychology simply do not understand the concepts of reliability and validity as applied to measurement -even some graduates of Stanford- believe it or not.

Forgive me for returning to the Taba project for a specific example. During our attempt to validate our test of 'ability to apply ideas,' we interviewed some 31 sixth graders. The format required the reading of a fictitious scenario followed by a series of statements to which the student answered `probably true' or `probably false', and why. The scenario in this instance describes two groups of people living in the same valley, one a hunting society, the other a farming society. One of the items (subsequently revised) was: the Hunteros live the way they do because they are less able to learn than the Farmanos.: The ideal response would be something like the following: "Probably false because both groups learn what they need to -you can't tell about ability to learn from how they live." Three students gave essentially that response.

Examples of other responses are as follows: "Well, I think that's false because -well, they move around a lot and they can learn things." Apparently, this student focused on the opportunities provided to the hunters for learning, rather than relative 'ability to learn.' Subsequent questioning would have been necessary to determine whether in fact, the student understood our intended idea.

Another example: "That's probably false because neither one of them have schools." This response, strongly suggests the "misconception" that learning only occurs in school as well as the "incorrect" inference that neither group would have schools. One also wonders about the student's concept of school.

A more common response was "Probably true because only Farmanos will have schools."

One final example: "That's probably true because they have the tools to do all their fishing, hunting, cooking, and stuff." This response suggests that the student confused ability to learn with need to learn and infered that the hunters have no need to learn.

I wish I could report that subsequent questioning clarified these interpretations and that post-Taba all students clearly comprehended the important idea that life style is no index to ability to learn. Unfortunately neither is the case, we simply don't know, and that is my point.

For a second illustration, I have chosen a recent research report which appeared in Social Education. Before reviewing this report, I should make it clear that I am not judging it to be a poor piece of work. On the contrary, I think it is a good study within the current research framework. It is the latter which I believe to be misguided.

First, on the positive side, the study intended to "explore the qualities of student reasoning about basic economic concepts and to observe any changes in student reasoning across grade levels." The seven concepts included five which had been previously identified as fundamental to understanding other economic ideas. Examples are "economic wants" and "opportunity cost." The method used to assess reasoning consisted of a structured interview focused around specific situations and questions which, from the examples given, seem promising as ways to assess grasp of the concepts.

So far, so good. But, lo and behold we are to discover nothing about the accuracy or appropriate use of these concepts. Instead, responses were categorized as unreflective, transitional or reflective a la Piaget. Responses were considered reflective if they involved forming hypotheses beyond immediate experience or included value statements or set criteria for decision making or "involved ideas about broader community concerns." Not surprisingly, all concepts showed increase in reflective thinking from grades 1 to 9 albeit at differing rates for different concepts. The author concludes that reasoning about economic concepts fits cognitive development theory in that it becomes more abstract, other directed and flexible with increasing age. This, we are told, suggests that educators need to use the findings of theorists such as Piaget to assist in developing a curriculum appropriate for students at each grade level.

But how is this to be done? The results suggest that fourth graders, for example, are very "reflective" in their reasoning about "opportunity cost," but not about "economic wants." It may be that this indicates poorer comprehension of the latter concept but it may not indicate this at all. Given the research strategy, an equally good interpretation is simply that kids can use more big words and express more opinions with regard to the materials subsumed under the "cost" concept. Isn't it time after twenty years to question whether Piaget really has anything to say to curriculum builders?

And there are other problems. Consider that the sample studied consisted of 72 students, 24 in each of grades 1, 4, and 9, in an unidentified suburban community and chosen for high socioeconomic status and cognitive skill because "the small sample size required that the sample be homogeneous." How are we to base curriculum on data from such a sample? Further, the article reflects no compunction whatsoever about using

multivariate Analysis of Variance, followed by univariate tests for each concept, not only for linear trend, but for curvilinearity as well. Just to be safe, Kendalls Tau was also calculated. What do these statistics contribute? Certainly there was no approximation even, of random sampling, nor is it easy to conjure up a meaningful hypothetical population from which the sample might be argued to approximate randomness. Inspection of the graphs is sufficient to conclude that for these particular students, all concepts, except one, showed increased "reflective reasoning" with age and that the pattern differed across concepts.

Most serious is the possible misconception suggested by the author's conclusion that "reflective economic reasoning was well developed for four out of five of these key concepts at grade four."

Maybe, but much of that reflective reasoning may be very confused reasoning as suggested by two examples of reflective thinking provided with respect to the concept of "monetary value." The student was presented with a U. S. one-dollar bill and an obvious play money dollar and asked which they preferred to have and why one was of value and the other was not. Two examples of reflective responses are: "the U. S. mint makes it" and "It is the money of the U. S. goverment." I submit that these reponses are a far cry from that of another student who said "money has value because we all agree on it." The latter, I think, represents the concept accurately.

I think this author has made the same mistakes we made on the Taba project, to wit: 1) looking too much to other disciplines for guidance in studying the educational process, 2) trying to study underlying psychological processes instead of important content, and 3) unnecessarily and somewhat inappropriately over-statisticizing.

As an alternative, I suggest identifying important concepts and ideas, preferably few in number -I seriously doubt that there are really that many that can be defended as necessary or even useful- and then finding out how well students understand them (we might even try using the taxonomy of educational objectives) and whether, after our best efforts, they understand them any better. This would require serious attention to the difficult task of assessing understanding, would certainly require true cooperation between teachers and researchers and would threaten much of our current research establishment. If you recall hearing much this same pitch being made by Hilda Taba -you're right, and so was she.